

## Eine Skizze zur Automatisierung qualitativer Inhaltsanalyse

Hoxtell, Annette

Veröffentlichungsversion / Published Version  
Zeitschriftenartikel / journal article

### Empfohlene Zitierung / Suggested Citation:

Hoxtell, A. (2019). Eine Skizze zur Automatisierung qualitativer Inhaltsanalyse. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(3). <https://doi.org/10.17169/fqs-20.3.3340>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:  
<https://creativecommons.org/licenses/by/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:  
<https://creativecommons.org/licenses/by/4.0>

## Eine Skizze zur Automatisierung qualitativer Inhaltsanalyse

Annette Hoxtell

### Keywords:

qualitative  
Inhaltsanalyse;  
Automatisierung;  
Teilautoma-  
tisierung;  
Digitalisierung;  
künstliche  
Intelligenz;  
Kategorien;  
Kategoriensystem;  
Kategorisierung;  
Text Mining;  
Transkription

**Zusammenfassung:** Die (Teil-) Automatisierung qualitativer Inhaltsanalyse ist absehbar und technisch bereits in Teilen möglich. Ich skizziere anhand einer von mir durchgeführten qualitativen Inhaltsanalyse eine fiktive Automatisierungslösung, die auch die Datenerhebung und Transkription umfasst. Es handelt sich hierbei um eine Teilautomatisierung, da sie die kontinuierliche Zusammenarbeit von Mensch und Maschine voraussetzt. Wo bereits existent, zeige ich Automatisierungsansätze auf, die sich für die Integration in eine (teil-) automatisierte qualitative Inhaltsanalyse eignen.

### Inhaltsverzeichnis

- [1. Einführung](#)
- [2. Die initial durchgeführte qualitative Inhaltsanalyse](#)
- [3. Qualitative Inhaltsanalyse und Automatisierung](#)
- [4. \(Teil-\) Automatisierung Qualitativer Inhaltsanalyse](#)
  - [4.1 Vorarbeit: Transkription](#)
  - [4.2 Hauptarbeit: qualitative Inhaltsanalyse](#)
  - [4.3 Erweiterung: Datenerhebung und kommunikative Validierung](#)
- [5. Fazit und Ausblick](#)

[Literatur](#)

[Zur Autorin](#)

[Zitation](#)

## 1. Einführung

2014 und 2015 interviewte ich 72 Personen für meine Dissertation, transkribierte die Interviews und unterzog sie einer qualitativen Inhaltsanalyse. Monatelang verlor ich mich in ihren Tiefen: Repräsentierte das Kategoriensystem alle inhaltstragenden Textstellen gebührend? War die Einteilung in Ober- und Unterkategorien sinnvoll? Wie sollte ich mit Textstellen verfahren, die mehreren Kategorien zugeordnet werden konnten? Häufig hätte ich diesen repetitiven und zeitaufwändigen Prozess gerne abgekürzt, manchmal ihn am liebsten komplett abgegeben. Die qualitative Inhaltsanalyse müsste sich doch automatisieren lassen, dachte ich mir. In den Folgejahren, in denen ich weitere qualitative Inhaltsanalysen – jedoch geringeren Umfangs – durchführte und Studierende mit dem Verfahren vertraut machte, ließ mich diese Idee nicht mehr los. [1]

Im Folgenden stelle ich kurz das Forschungsdesign und die von mir initial durchgeführte qualitative Inhaltsanalyse dar (Abschnitt 2), kläre wichtige Begriffe (Abschnitt 3) und zeige dann auf, wie eine mögliche Teilautomatisierung des Verfahrens aussehen könnte und welche Ansätze bereits existieren (Abschnitt 4). Abschließend gehe ich darauf ein, für wen und unter welchen Bedingungen sich eine solche Teilautomatisierung anböte (Abschnitt 5). [2]

## 2. Die initial durchgeführte qualitative Inhaltsanalyse

Für meine Dissertation wählte ich ein exploratives Design, bei dem ich mich pragmatisch am methodischen Werkzeugkasten (STAMANN, JANSSEN & SCHREIER 2016) bediente. Die forschungsleitende Frage lautete: Aus welchen Gründen entscheiden sich Jugendliche für einen Ausbildungsbetrieb? Narrative Interviews, bei denen die Erfahrungen, Ansichten und Motive der Beforschten im Mittelpunkt standen (CRESWELL 2007; SCHÜTZE 1976; WATTANASUWAN, BUBER & MEYER 2007), bildeten das Kernstück. Ein Teil dieser Interviews war in vier multiple Fallstudien eingebettet (YIN 2009), die zwecks Daten- und Methodentriangulation (FLICK 2011) zusätzliche Expert\_inneninterviews (BORTZ & DÖRING 2016; PFADENHAUER 2009), teilnehmende Beobachtungen und Werbemedien miteinschlossen. [3]

Sämtliche Interviews transkribierte ich in die literarische Umschrift (MAYRING 2016). Die narrativen und Expert\_inneninterviews unterzog ich einer strukturierenden Inhaltsanalyse (KUCKARTZ 2016) in Anlehnung an MAYRINGS (2015) Zusammenfassung und induktive Kategorienbildung sowie Strukturierung/deduktive Kategorienanwendung, da die Verfahren stark regelgeleitet sind und es mir ermöglichten, Inhalte zu verdichten und thematisch zu ordnen (NADERER 2007; SCHREIER 2014). Dabei folgte ich STEIGLEDERS Modell, überarbeitete die Kategorien kontinuierlich (2007) und bezog die einzelnen Analyseschritte, wie in KUCKARTZ' (2016) Ablaufmodell der qualitativen Inhaltsanalyse, beständig auf die Forschungsfrage zurück. Abbildung 1 zeigt das konkrete Vorgehen.

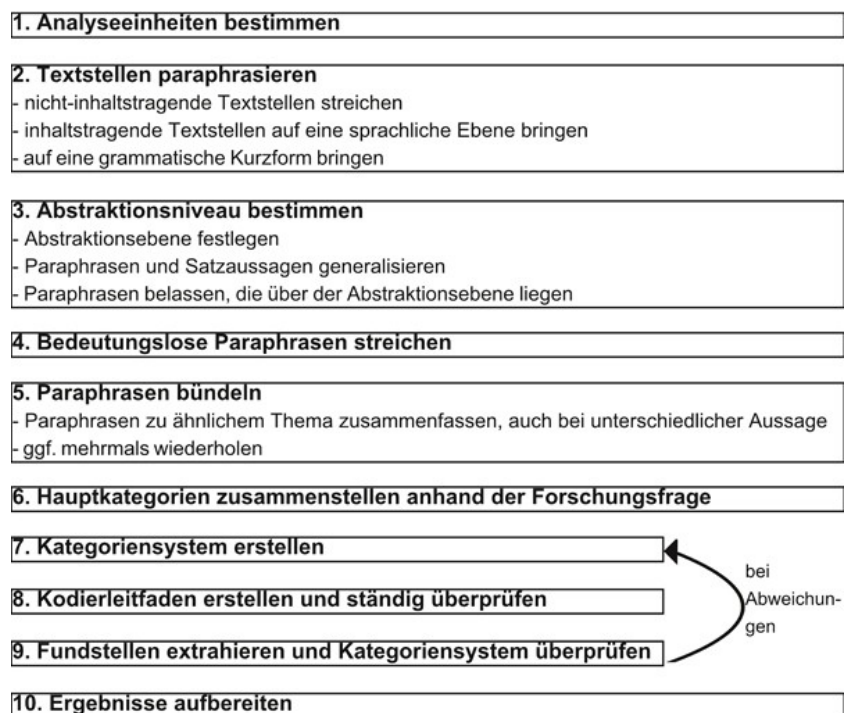


Abbildung 1: Ablauf der zusammenfassend-strukturierenden Inhaltsanalyse (HOXTELL 2016, S.42; angelehnt an MAYRING 2010, S.68-70; STEIGLEDER 2007, S.183) [4]

### 3. Qualitative Inhaltsanalyse und Automatisierung

Bevor ich auf Automatisierungsansätze der qualitativen Inhaltsanalyse eingehe, möchte ich die Begriffe *qualitative Inhaltsanalyse*, *Digitalisierung* und *Automatisierung* klären. Unter *qualitativer Inhaltsanalyse* verstehe ich ein regelgeleitetes, interpretatives Vorgehen, das an Validität und Reliabilität ausgerichtet wird und bei dem aus allen Daten systematisch Kategorien gebildet werden (KUCKARTZ 2016; NADERER 2007; SCHREIER 2014; STAMANN et al. 2016; STEIGLEDER 2007). Varianten der qualitativen Inhaltsanalyse werden von vielen Wissenschaftler\_innen entwickelt und genutzt; sie unterscheiden sich hinsichtlich des Grades an deduktiver und induktiver Kategorienbildung und hinsichtlich der Bezeichnung des Auswertungsverfahrens (BOYATZIS 1998; HSIEH & SHANNON 2005; KUCKARTZ 2016; MAYRING 2000, 2015; SCHREIER 2012, 2014). Typischerweise werden diese Daten in ihrer verschriftlichten Form analysiert. Qualitative Inhaltsanalysen zeichnet aus, dass Daten kontextabhängig interpretiert werden, also auch ihr Zustandekommen und ihre Wirkung berücksichtigt werden (MAYRING 2015). Auch können Mehrdeutigkeiten zugelassen werden (KUCKARTZ 2016). Damit unterscheiden sie sich von quantitativen Analysen, worunter auch die diktionsbasierte, computerunterstützte Inhaltsanalyse fällt. [5]

Voraussetzung für die Automatisierung der qualitativen Inhaltsanalyse ist die *Digitalisierung*. Digitalisierung bedeutet zum einen, analoge Inhalte in eine digitale Form zu übertragen und diese zum anderen von künstlichen Systemen bearbeiten zu lassen (WEBER & BUSCHBACHER 2017). Inhaltsanalysen werden heutzutage vielfach in digital unterstützter Form durchgeführt, beispielsweise wenn das Computerprogramm die Häufigkeitsverteilung von Kategorien in unterschiedlichen Datensätzen selbstständig aus zählt, so z.B. im Falle von [ATLAS.ti](#), [audiotranskription](#) und [MAXQDA](#). [6]

Einer kompletten *Automatisierung* der qualitativen Inhaltsanalyse entspricht dieser Einsatz digitaler Werkzeuge noch nicht. *Automatisierung* heißt, dass künstliche Systeme ehemals von Menschen ausgeübte Tätigkeiten übernehmen. Meist steuern sie Prozesse und treffen Entscheidungen (VOIGT o.D.), wie z.B. Sensoren, die den Öl- oder Gasdurchfluss in Pipelines mittels Ultraschall überwachen und Lecks oder Verstopfungen melden (GEIPEL-KERN 2011), ehe diese für das menschliche Auge sichtbar werden. Eine automatisierte qualitative Inhaltsanalyse müsste neue Kategorien bilden, wenn das Datenmaterial es erfordert.<sup>1</sup> [7]

Da sie dies selbständig täte und ihr Handeln an sich ändernde Umstände anpasste, anstatt nur Vorgaben abzuarbeiten, würde es sich bei einer automatisierten qualitativen Inhaltsanalyse um eine schwache Form *künstlicher Intelligenz* handeln (NATIONAL SCIENCE AND TECHNOLOGY COUNCIL-COMMITTEE ON TECHNOLOGY 2016; PEINL 2018a; SCHMIECH 2018). [8]

---

<sup>1</sup> Die Automatisierung übernimmt in dieser Gedankenskizze menschliche Tätigkeiten. Aus diesem Grund werden Anthropomorphismen bewusst verwendet.

## 4. (Teil-) Automatisierung Qualitativer Inhaltsanalyse

### 4.1 Vorarbeit: Transkription

Bereits die Transkription, oft der vorbereitende Schritt einer qualitativen Inhaltsanalyse, ärgerte mich: Wieso sollte ich 72 Interviews abtippen? Oder jemanden dafür bezahlen? Wieso ging das nicht automatisiert? Regelmäßig auf der Mailingliste Qualitative Sozialforschung (QSF-L) gestellte Fragen zeigen, dass andere Sozialforschende ähnlich denken. So wurde z.B. am 23.11.2017 gefragt: "Gibt es unter Euch/Ihnen Erfahrung mit direkter Spracherkennung im Zusammenhang mit Interviews? Oder bleibt die nachträgliche Transkription der Königsweg".<sup>2</sup> [9]

Ich wünsche mir ein System, das erkennt, welche Art von Daten vorliegt, egal ob in Audio-, Video- oder Textform (GÖBEL 2018). Es soll anhand des Klangbildes identifizieren, wer spricht (BITKOM 2018) und das Gesprochene transkribieren (a.a.O.). Aktuell sieht es so aus: Künstliche Systeme erkennen recht zuverlässig, welche Arten von Daten vorliegen und ob sich Muster in ihnen verbergen, beispielsweise bei Spam-Filtern, der Foto- und Handschriftenerkennung oder selbstfahrenden Autos (CLARK 2016; FEHRENBACHER 2015; LÜTTERS 2017; PAKALSKI 2006). Für die qualitative Inhaltsanalyse interessieren vor allem sprachgesteuerte Assistenzsysteme wie Alexa, Bixby, Cortana, Google Home, Siri und Watson. Sie unterscheiden menschliche Stimmen von Umgebungsgeräuschen, reagieren auf Aufforderungen (LÜTTERS 2018) und lernen daraus. [10]

Speziell für die Transkription eignet sich Diktiersoftware wie [Dragon](#) oder das in Office integrierte [Diktat](#). Allerdings sind sie nur auf eine Person ausgelegt. Ab Sommer 2019 soll die Software [f4transkript](#) in der Lage sein, Gespräche mit zwei oder mehr Teilnehmenden in Text umzuwandeln. Dabei soll der Zeitaufwand für die Korrektur bei zwei Sprechenden erstmals geringer ausfallen als der für die händische Transkription. [11]

### 4.2 Hauptarbeit: qualitative Inhaltsanalyse

Zuerst las ich einen Teil der Interviews als Ganzes, notierte, was mir spontan dazu einfiel und unterstrich Textstellen, die mir für die Kategorisierung relevant erschienen (KUCKARTZ 2016). Durch diese Unterstreichungen legte ich Stück für Stück die *Analyseeinheiten* fest. Schon an dieser Stelle hätte ich mir ein System gewünscht, das dieselben Arbeitsschritte parallel ausführt. So hätte ich meine mit seinen Notizen und Unterstreichungen vergleichen und im Anschluss das *Abstraktionsniveau* bestimmen und *Textstellen so paraphrasieren* können, dass sie dem festgelegten Abstraktionsniveau entsprechen. Dem System hätte ich mitgeteilt, mit welchen von ihm vorgeschlagenen Analyseeinheiten ich einverstanden bin und mit welchen nicht. Für einen Teil der akzeptierten

2 Die Mail ist für Abonnent\_innen von QSF-L unter [https://lists.fu-berlin.de/private/qsf\\_l/2017-November/msg00041.html](https://lists.fu-berlin.de/private/qsf_l/2017-November/msg00041.html) zugänglich. Eine Registrierung ist über [https://lists.fu-berlin.de/listinfo/qsf\\_l](https://lists.fu-berlin.de/listinfo/qsf_l) möglich.

Analyseeinheiten hätte ich ihm Paraphrasen vorgegeben; die restlichen Analyseeinheiten hätte es selbständig paraphrasiert. Im nächsten Schritt hätte es *Paraphrasen gebündelt* und solche *ohne Bedeutung gestrichen*. Sein Vorgehen hätte es ähnlich dem nicht-automatisierten Vorgehen in Tabelle 1 dokumentiert.

Fall, Zeile	Textstelle	Paraphrase	Bündelung
1, 12	vertrauenswürdige Menschen, die mich akzeptieren, wie ich bin	Bei der Arbeit als ganzer Mensch akzeptiert werden	K 4 Personal Umgang miteinander Menschenbild
3, 25	Ich glaube eher, also ich bin eher ein Mensch, der freundschaftlich mit Kollegen umgehen würde, weil ich mag es nicht, wenn man so tut, als ob man sich kaum kennt, obwohl man ja zusammenarbeitet und das ist, glaube ich, erleichtert viel, wenn man sich versteht	Freundschaftlicher Umgang	
7, 29	Die haben auch Interesse an mir gezeigt. Die wollten was über mich erfahren. Was ich mache in meiner Freizeit. Wie ich darauf gekommen bin auf diesen Beruf.	Bei der Arbeit als ganzer Mensch akzeptiert und gesehen werden	

Tabelle 1: Paraphrasierung und Bündelung [12]

Hier wäre wieder ein Kontrollschritt notwendig geworden: Ähnlich der Überprüfen-Funktion in Word hätte ich Paraphrasen und Bündelungen angenommen oder diese abgelehnt und ersetzt. Indem das System den gesamten Kodierprozess (BITKOM 2018) einschließlich der Entscheidungsbegründungen dokumentiert hätte, hätte es seine Kodierentscheidungen nachvollziehbar und transparent gestaltet (DIE BUNDESREGIERUNG 2018). [13]

Anhand der Forschungsfrage und der Bündelungen *stellte ich Hauptkategorien zusammen* und *erstellte ein Kategoriensystem*. Der Ausschnitt in Tabelle 1 gehört

beispielsweise zur Hauptkategorie "Ausbildungsbetrieb" und zur Unterkategorie "Personal" bei Schüler\_innen. Dieses Kategoriensystem bildete ich in einem *Kodierleitfaden* ab; ein Ausschnitt findet sich in Tabelle 2.

Kategorie	Definition	Ankerbeispiel	Regeln
1.3.1 Gründe für Betrieb	Gründe für Ausbildungsbetrieb, die breit gefasst sind	Dass ich mich wohlfühle. Also es hängt alles zusammen. Das Gesamte muss stimmen.	Gründe, die nicht in spezifische Unterkategorie fallen
1.3.1.1 Unternehmensklima	Arbeitsumfeld, Art des Miteinanders im Betrieb	sozial, nicht zu spießig, zuvorkommend	Das kann sowohl eine Beschreibung der Betriebseigenschaften sein, als auch dessen, was sie mit dem/der Schüler_in machen
1.3.1.2 Verdienst	Gehalt	Einmal, äh, Geld ist doch wichtig.	Was sie verdienen wollen
1.3.1.3 Spaß	Spaß/Freude an Arbeit	Ich möchte schon Spaß auf Arbeit haben.	Die intrinsische Motivation der Auszubildenden: Spaß an der Tätigkeit und am Miteinander

Tabelle 2: Kodierleitfaden [14]

Auch hier hätte ich mir ein System gewünscht, das die gleichen Arbeitsschritte für die ersten Interviews parallel ausführt und es mir erlaubt, seinen und meinen Kodierleitfaden miteinander zu vergleichen sowie von ihm vorgeschlagene Kategorien und ihre Beziehungen zueinander anzunehmen oder zu korrigieren. [15]

Basierend auf dem von mir überprüften und angepassten initialen Kategoriensystem hätte das System selbständig Fundstellen in den restlichen Interviews extrahiert sowie den Kodierleitfaden und das Kategoriensystem beständig überprüft und angepasst. Bei Unklarheiten, beispielsweise mehreren Kategorisierungsmöglichkeiten, hätte es die Analyse gestoppt und mich um eine Entscheidung gebeten (GÖBEL 2018; WEBER & BUSCHBACHER 2017). Wichtig wäre, dass ich jede Stelle des Kodierprozesses ansteuern und Entscheidungen des künstlichen Systems rückgängig machen könnte. [16]

Aktuell sieht es so aus: Eine komplette Automatisierung der qualitativen Inhaltsanalyse gibt es nicht. Jedoch existieren Verfahren, die zur Festlegung der Analyseeinheiten, zur Bildung von Hauptkategorien sowie zu einer einfachen Bedeutungseinschätzung genutzt werden können. Analyseeinheiten lassen sich



mittels *Text Mining* bestimmen. Dieses Verfahren wird bereits genutzt, z.B. um Texte in bibliografischen Datenbanken wie MEDLINE zu erschließen (PAPANIKOLAOU, PAVLOPOULOS, THEODOSIOU & ILIOPOULOS 2015; WESTERGAARD, STÆRFELDT, TØNSBERG, JENSEN & BRUNAK 2018); es soll auch bei der Deutschen Nationalbibliothek Anwendung finden (JUNGER & SCHWENS 2017). Hauptkategorien können künstliche Systeme bereits zusammenstellen. Im Online-Buchhandel werden Genres automatisch gebildet (WEBER & BUSCHBACHER 2017), und bei der Auswertung des "Bürgerdialogs" der Bundesregierung bedienten sich Sozialforscher\_innen automatisierter Frequenz- und Differenzanalysen, um Oberkategorien aus Online-Textbeiträgen zu erstellen. Manuell validierten sie diese und ergänzten sie um Unterkategorien (WALDHERR et al. 2019). Diese manuelle Validierung entspricht den von mir im gesamten Prozess geforderten Kontrollschritten und der Zusammenarbeit von Mensch und Maschine, die auch in anderen Anwendungsfeldern künstlicher Intelligenz genutzt wird. In der Krebsvorsorge z.B. holen sich Ärzt\_innen eine automatisierte Zweitmeinung ein und kommen gemeinsam zu einer besseren Diagnose (WANG, KHOSLA, GARGEYA, IRSHAD & BECK 2016), und die amerikanische Fallrechtsuchmaschine Judicata zieht Menschen bei Unklarheiten zur Hilfe hinzu und legt ihre Entscheidungen offen (GURARI 2017). Der übersichtsartigen Einschätzung von Beiträgen dient die *Sentiment-Analyse*; sie wird z.B. genutzt, um negative von positiven und neutralen Kund\_innenkommentaren zu trennen (BANNISTER 2018; KREUTZER & SIRRENBORG 2019). Die qualitative Datenanalyse (QDA)-Software "Texifter" ermöglicht eine Teilautomatisierung der Kategorienbildung sowie der Bedeutungseinschätzung; menschliche Vorarbeit und Validierung werden dabei zwingend vorausgesetzt (PRIEBE 2016; R. 2019; siehe auch z.B. [DiscoverText](#)). [17]

#### 4.3 Erweiterung: Datenerhebung und kommunikative Validierung

Ich wünsche mir, neben der Transkription und der qualitativen Inhaltsanalyse, weiterhin bestimmte Formen der Datenerhebung zu automatisieren. Anstatt mit allen Schüler\_innen einzeln zu sprechen, hätten spätere Interviews automatisiert durchgeführt werden können. In einem ersten Befragungsschritt hätten die Teilnehmenden offene Fragen beantwortet; die Antworten hätte das System zeitgleich einer qualitativen Inhaltsanalyse unterzogen. In einem zweiten Schritt hätte das System die Schüler\_innen mit Antwortmöglichkeiten konfrontiert, die von anderen Teilnehmenden genannt worden waren. So hätten sich mit einer Erhebung sowohl neue Antworten generieren als auch eine Quantifizierung des gesamten Antwortgeschehens erreichen lassen. Anstatt den Umweg über die Schriftsprache zu gehen, könnte das künstliche System auch mündliche Interviews durchführen, analysieren und die Ergebnisse im Audioformat ausgeben. [18]

In der Marktforschung wird genau diese automatisierte Sprachein- und -ausgabe mithilfe von Sprachassistenten getestet, wie in Abschnitt 4.1 als Ersatz der herkömmlich-händischen Transkription vorgeschlagen. Befragte können Aufnahmen noch einmal anhören, das Transkript gegenlesen und es je nach Einstellung anpassen (LÜTTERS 2017). Damit fordert das System automatisch



zur kommunikativen Validierung auf (FLICK 2002) – auditiv oder visuell, je nach Vorliebe der Interviewten – und trägt somit zur Erhöhung der internen Validität bei. [19]

In der folgenden Tabelle 3 gebe ich einen Überblick über die in den vorherigen Abschnitten vorgeschlagene Teilautomatisierung der qualitativen Inhaltsanalyse. Dieses Vorgehen ließe sich auch an andere Varianten der qualitativen Inhaltsanalyse anpassen.

<b>Arbeitsschritte</b>	<b>Mensch</b>	<b>Maschine</b>
Vorarbeit: Transkription	- kontrolliert Transkripte und passt ggf. an	- erkennt den Datentyp - erkennt, wer spricht - transkribiert
Analyseeinheiten festlegen	- für alle Texte - vergleicht eigene mit maschinellen Analyseeinheiten und passt ggf. an	- für alle Texte
Paraphrasieren + Abstraktionsniveau bestimmen + Bündelung + Streichung von Paraphrasen	- macht dies für einen Teil der Texte - kontrolliert und passt ggf. an	- macht dies für den Rest der Texte
Hauptkategorien zusammenstellen + Kategoriensystem erstellen + Kodierleitfaden erstellen	- macht dies anhand eines Teils der Texte - greift bei Unklarheiten ein - kontrolliert Kodier- und Kategorisierungsprozess und passt ggf. an	- macht dies für den Rest der Texte - überarbeitet Kodierleitfaden und Kategoriensystem kontinuierlich - stoppt Analyse bei Unklarheiten und bittet um menschlichen Input - dokumentiert Vorgehen
Ergebnisaufbereitung	- macht dies komplett	

Arbeitsschritte	Mensch	Maschine
Erweiterung: Datenerhebung + Auswertung	<ul style="list-style-type: none"> <li>- führt die ersten Interviews</li> <li>- kontrolliert und passt ggf. an</li> </ul>	<ul style="list-style-type: none"> <li>- führt weitere Interviews</li> <li>- führt qualitative Inhaltsanalyse in Echtzeit durch</li> <li>- stellt geschlossene Fragen basierend auf Antworten aus früheren Interviews</li> <li>- wertet Antworten quantitativ aus</li> <li>- fordert zur kommunikativen Validierung auf</li> </ul>

Tabelle 3: Skizze einer teilautomatisierten qualitativen Inhaltsanalyse [20]

## 5. Fazit und Ausblick

Gerade Personen, die alleine eine qualitative Inhaltsanalyse durchführen und nicht Teil größerer Forschungsprojekte mit mehreren Codierer\_innen sind, können von einer Teilautomatisierung der qualitativen Inhaltsanalyse profitieren. Ein künstliches System kann zur Reflexion der eigenen Kategorisierung anregen. Im Gegensatz zu Teilnehmer\_innen an Forschungswerkstätten und Interpretationsgruppen kann es das gesamte Datenmaterial und nicht nur Teile davon analysieren. [21]

Künstliche Systeme benötigen sehr viel Trainingsmaterial zum Lernen (PEINL 2018a); insofern böte sich eine Teilautomatisierung der qualitativen Inhaltsanalyse gerade auch für größere Forschungsvorhaben an, bei denen eine entsprechend große Datenbasis vorliegt, die nach einheitlichen Regeln kodiert wird. Unabhängig von der Größe des Forschungsvorhabens würde eine *Open-Data*-Strategie allen Forschenden bei der Teilautomatisierung der qualitativen Inhaltsanalyse nutzen. Große Mengen qualitativ hochwertiger Daten aus Wissenschaft und öffentlichem Dienst würden, wie von der BUNDESREGIERUNG (2018) gefordert, dezentral rechtskonform verfügbar gemacht. Offene Schnittstellen ermöglichten in diesem Fall die zweckgebundene Datenübermittlung (WEBER & BUSCHBACHER 2017). Monopolen und der Machtkonzentration der datensammelnden und verarbeitenden Stellen würde so entgegengewirkt (PEINL 2018b). Zum einen könnten der Kodierprozess und -leitfaden offengelegt werden (SCHARKOW 2012), zum anderen auch die Fundstellen aus früheren Projekten. Diese könnten ähnlich studentischen Arbeiten in der Plagiaterkennungssoftware [PlagScan](#) anonymisiert angezeigt werden; bei Nachfragen bestünde die Möglichkeit, diejenigen zu kontaktieren, die sie eingestellt haben. Eine solche datenschutzkonforme Offenlegung sensibler Daten geschieht z.B. im Medical Data Space der Fraunhofer-Gesellschaften. Dort

werden Gesundheitsdaten dezentral verwaltet und verarbeitet und bei Bedarf ausgetauscht (WEBER & BUSCHBACHER 2017, siehe auch den [Fraunhofer Medical Data Space](#)). Die 2016 ins Leben gerufene [Rotterdam Exchange Format Initiative](#) schafft eine Grundlage für den Austausch von Daten, indem sie einen Interoperabilitätsstandard festlegt. Mit ihm können qualitative Daten mit verschiedenen QDA-Programmen bearbeitet und somit z.B. auch einer qualitativen Inhaltsanalyse unterzogen werden. [22]

Eine teilweise Automatisierung der qualitativen Inhaltsanalyse halte ich im Sinne einer Mensch-Maschine-Interaktion für sinnvoll. Auch existiert sie bereits in Teilen, die jedoch zu einem teilautomatisierten System zusammengefügt und weiterentwickelt werden müsste. Der Mensch ist dem künstlichen System in der hermeneutisch-interpretativen Leistung überlegen, das System besser darin, große Datenmengen zu untersuchen. Voraussetzung für die Teilautomatisierung ist, dass Forschende das Auswertungsgeschehen nachvollziehen und in es eingreifen können. Aufgrund dieses Eingreifens und der Arbeitsteilung kann nur von einer Teilautomatisierung die Rede sein; eine Vollautomatisierung ist zum jetzigen Zeitpunkt nicht denkbar und würde auch dem Anspruch an eine qualitative Analyse nicht gerecht. [23]

## Literatur

- Bannister, Kristian (2018). Understanding sentiment analysis: What it is & why it's used. *Blogbeitrag*, 26. Februar, <https://www.brandwatch.com/blog/understanding-sentiment-analysis/> [Zugriff: 21. Juni 2019].
- BITKOM (2018). *Digitalisierung gestalten mit dem Periodensystem der Künstlichen Intelligenz*. Berlin: Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V., [https://bitkom.de/sites/default/files/2018-12/181204\\_LF\\_Periodensystem\\_online\\_0.pdf](https://bitkom.de/sites/default/files/2018-12/181204_LF_Periodensystem_online_0.pdf) [Zugriff: 21. Januar 2019].
- Bortz, Jürgen & Döring, Nicola (2016). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Boyatzis, Richard E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage.
- Clark, Andrew (2016). Machine Learning mit der Crowd bietet viele Vorteile. *Blogbeitrag*, 30. Oktober, <https://www.crowdguru.de/blog/machine-learning-mittels-crowd/> [Zugriff: 20. April 2018].
- Creswell, John W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2. Aufl.). Thousand Oaks, CA: Sage.
- Die Bundesregierung (2018). *Eckpunkte der Bundesregierung für eine Strategie Künstliche Intelligenz*, [https://www.bmbf.de/files/180718%20Eckpunkte\\_KI-Strategie%20final%20Layout.pdf](https://www.bmbf.de/files/180718%20Eckpunkte_KI-Strategie%20final%20Layout.pdf) [Zugriff: 22. Januar 2019].
- Fehrenbacher, Katie (2015). How Tesla is ushering in the age of the learning car. *Fortune*, 16. Oktober, <http://fortune.com/2015/10/16/how-tesla-autopilot-learns/> [Zugriff: 20. April 2018].
- Flick, Uwe (2002). *Qualitative Sozialforschung. Eine Einführung*. Hamburg: Rowohlt.
- Flick, Uwe (2011). *Triangulation: eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Geipel-Kern, Anke (2011). Bilder: Zehn Beispiele aus der Praxis – Wie Automatisierung die Energieeffizienz verbessert. *Process*, 9. September, <https://www.process.vogel.de/index.cfm?pid=7513&pk=333136&fk=0&type=article#2> [Zugriff: 29. Januar 2019].
- Göbel, Richard (2018). Big Data – Herausforderung bei der Nutzung großer Datenmengen. In Dietmar Wolff & Richard Göbel (Hrsg.), *Digitalisierung: Segen oder Fluch? Wie die Digitalisierung unsere Lebens- und Arbeitswelt verändert* (S.73-99). Berlin: Springer.

- Gurari, Itai (2017). Embedding humans in the loop. *judicata Blogbeitrag* 17. November, <https://blog.judicata.com/embedding-humans-in-the-loop-2c19fb41c926> [Zugriff: 28. Januar 2019].
- Hoxtell, Annette (2016). Entscheidungsgründe für den Ausbildungsbetrieb – eine explorative Studie mit Auszubildenden sowie Schülerinnen und Schülern. *Dissertation*, Wirtschaft Arbeit Technik, Universität Potsdam, <http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-93377> [Zugriff: 13. Juni 2019].
- Hsieh, Hsiu-Fang & Shannon, Sarah E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288. <https://doi.org/10.1177/1049732305276687> [Zugriff: 21. Januar 2019].
- Junger, Ulrike & Schwens, Ute (2017). Die inhaltliche Erschließung des schriftlichen kulturellen Erbes auf dem Weg in die Zukunft. *DNB professionell*, [https://www.dnb.de/SharedDocs/Downloads/DE/Professionell/Erschliessen/inhaltlicheErschliessungZukunft.pdf?\\_\\_blob=publicationFile&v=2](https://www.dnb.de/SharedDocs/Downloads/DE/Professionell/Erschliessen/inhaltlicheErschliessungZukunft.pdf?__blob=publicationFile&v=2) [Zugriff: 12. Juni 2019].
- Kreutzer, Ralf & Sirrenberg, Marie (2019). *Künstliche Intelligenz verstehen: Grundlagen – Use-Cases – unternehmenseigene KI-Journey*. Wiesbaden: Springer.
- Kuckartz, Udo (2016). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (3. Aufl.). Weinheim: Beltz Juventa.
- Lütters, Holger (2017). Sprich mit mir!. *planung & analyse*, 17(5), 58-60.
- Lütters, Holger (2018). Digitale Sprachassistent: in der Forschung. Von der Technologie zur Anwendung. In Matthias Knaut (Hrsg.), *KREATIVITÄT + X = INNOVATION Beiträge und Positionen der HTW Berlin* (S.34-41). Berlin: BWV.
- Mayring, Philipp (2000). Qualitative Inhaltsanalyse. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(2), Art. 20, <https://doi.org/10.17169/fqs-1.2.1089> [Zugriff: 4. August 2019].
- Mayring, Philipp (2010). *Qualitative Inhaltsanalyse Grundlagen und Techniken* (11. Aufl.). Weinheim: Beltz.
- Mayring, Philipp (2015). *Qualitative Inhaltsanalyse Grundlagen und Techniken* (12. Aufl.). Weinheim: Beltz.
- Mayring, Philipp (2016). *Einführung in die qualitative Sozialforschung: eine Anleitung zu qualitativem Denken* (6. überarb. Aufl.). Weinheim: Beltz.
- Naderer, Gabriele (2007). Auswertung & Analyse qualitativer Daten. In Gabriele Naderer & Eva Balzer (Hrsg.), *Qualitative Marktforschung in Theorie und Praxis: Grundlagen, Methoden und Anwendungen* (S.363-391). Wiesbaden: Gabler.
- National Science and Technology Council, Committee on Technology (2016). *Preparing for the future of artificial intelligence*. Washington, D.C.: Executive Office of the President. <https://info.publicintelligence.net/WhiteHouse-ArtificialIntelligencePreparations.pdf> [Zugriff: 23. Januar 2019].
- Pakalski, Ingo (2006). Strato: Filter erkennt Bilder-Spam. *golem.de*, 14. Dezember, <https://www.golem.de/0612/49471.html> [Zugriff: 20. April 2018].
- Papanikolaou, Nikolas; Pavlopoulos, Georgios A.; Theodosiou, Theodos & Iliopoulos, Ioannis (2015). Protein-protein interaction predictions using text mining methods. *Methods*, 74, 47-53.
- Peinl, René (2018a). Privatleben 4.0 – Wie Digitalisierung, das Internet der Dinge und Deep Learning unser Privatleben verändern. In Dietmar Wolff & Richard Göbel (Hrsg.), *Digitalisierung: Segen oder Fluch? Wie die Digitalisierung unsere Lebens- und Arbeitswelt verändert* (S.225-252). Berlin: Springer.
- Peinl, René (2018b). Smart Speaker ohne Cloud. *Video*, 14. Dezember, <https://www.youtube.com/watch?v=3UXzxghpcuo> [Zugriff: 23. Januar 2019].
- Pfadenhauer, Michaela (2009). Das Experteninterview: Ein Gespräch auf gleicher Augenhöhe. In Renate Buber & Hartmut. H. Holzmüller (Hrsg.), *Qualitative Marktforschung: Konzepte – Methoden – Analysen* (S.449-461). Wiesbaden: Gabler.
- Priebe, John (2016). Best practices for classifying uncoded data. *texifter*, 4. April, <https://texifter.zendesk.com/hc/en-us/articles/217965828-Best-practices-for-classifying-uncoded-data> [Zugriff: 21. Juni 2019].

- R., Iris (2019). Classify a large volume of data quickly. *texifter*, 6. Februar, <https://texifter.zendesk.com/hc/en-us/articles/202972680-Classify-a-large-volume-of-data-quickly> [Zugriff: 21. Juni 2019].
- Scharkow, Michael (2012). *Automatische Inhaltsanalyse und maschinelles Lernen*. Berlin: epubli.
- Schmiech, Chris (2018). Der Weg zur Industrie 4.0 für den Mittelstand – Ausgewählte Potenziale und Herausforderungen. In Dietmar Wolff & Richard Göbel (Hrsg.), *Digitalisierung: Segen oder Fluch? Wie die Digitalisierung unsere Lebens- und Arbeitswelt verändert* (S.1-28). Berlin: Springer.
- Schreier, Margrit (2012). *Qualitative content analysis in practice*. London: Sage.
- Schreier, Margrit (2014). Varianten qualitativer Inhaltsanalyse: ein Wegweiser im Dickicht der Begrifflichkeiten. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 15(1), Art. 18, <http://dx.doi.org/10.17169/fqs-15.1.2043> [Zugriff: 1. Februar 2016].
- Schütze, Fritz (1976). Zur Hervorlockung und Analyse von Erzählungen thematisch relevanter Geschichten im Rahmen soziologischer Feldforschung: dargestellt an einem Projekt zur Erforschung von kommunalen Machtstrukturen. In Ansgar Weymann & Arbeitsgruppe Bielefelder Soziologen (Hrsg.), *Kommunikative Sozialforschung: Alltagswissen und Alltagshandeln, Gemeindemachtforschung* (S. 159-260). München: Fink.
- Stamann, Christoph; Janssen, Markus & Schreier, Margrit (2016). Qualitative Inhaltsanalyse – Versuch einer Begriffsbestimmung und Systematisierung. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 17(3), Art. 16, <http://dx.doi.org/10.17169/fqs-17.3.2581> [Zugriff: 20. April 2018].
- Steigleder, Sandra (2007). *Die strukturierende qualitative Inhaltsanalyse im Praxistest: eine konstruktiv kritische Studie zur Auswertungsmethodik von Philipp Mayring*. Marburg: Tectum.
- Voigt, Kai-Ingo (o.D.). Definition: Automatisierung. *Gabler Wirtschaftslexikon*, <https://wirtschaftslexikon.gabler.de/definition/automatisierung-27138/version-250801> [Zugriff: 22. Januar 2019].
- Waldherr, Annie; Wehden, Lars-Ole; Stoltenberg, Daniel; Miltner, Peter; Ostner, Sophia & Pfetsch, Barbara (2019). Induktive Kategorienbildung in der Inhaltsanalyse: Kombination automatischer und manueller Verfahren. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(1), Art. 19, <http://dx.doi.org/10.17169/fqs-20.1.3058> [Zugriff: 29. Januar 2019].
- Wang, Dayong; Khosla, Aditya; Gargeya, Rishab; Irshad, Humayun & Beck, Andrew. H. (2016). Deep learning for identifying metastatic breast cancer. *arXiv:1606.05718* [cs, q-bio]. <https://arxiv.org/pdf/1606.05718.pdf> [Zugriff: 28. Januar 2019].
- Wattanasuwan, Kritsadarat; Buber, Renate & Meyer, Michael (2007). Das narrative Interview und die narrative Analyse. In Uwe Flick (Hrsg.), *Qualitative Sozialforschung: eine Einführung* (S.359-379). Hamburg: Rowolth.
- Weber, Mathias & Buschbacher, Florian (2017). *Künstliche Intelligenz – Wirtschaftliche Bedeutung, gesellschaftliche Herausforderungen, menschliche Verantwortung*. Berlin: BITKOM, DFKI, <https://www.bitkom.org/sites/default/files/file/import/171012-KI-Gipfelpapier-online.pdf> [Zugriff: 22. Januar 2019].
- Westergaard, David; Stærfeldt, Hans-Henrik; Tønsgaard, Christian; Jensen, Lars Juhl & Brunak, Søren (2018). A comprehensive and quantitative comparison of text-mining in 15 million full-text articles versus their corresponding abstracts. *PLOS Computational Biology*, 14(2), <https://doi.org/10.1371/journal.pcbi.1005962> [Zugriff: 13. Juni 2019].
- Yin, Robert K. (2009). *Case study research: Design and methods* (4. Aufl.). Los Angeles, CA: Sage.

## Zur Autorin

Annette HOXTELL ist Professorin für Allgemeine Betriebswirtschaftslehre an der Hochschule für Wirtschaft, Technik und Kultur Berlin. Ihre Forschungsinteressen gelten qualitativen Fragestellungen im Marketing und verwandten Bereichen. Sie lehrt Marketing, Interkulturelles und verwandte Fächer. Annette HOXTELL arbeitete als PR-Managerin und *Sustainability Advisor* in der IT-Branche und begeisterte als Leiterin eines Bildungsprojektes Mädchen für MINT-Berufswege.

Kontakt:

Prof. Dr. Annette Hoxtell

hwtk Berlin

Bernburger Straße 24/25, 10965 Berlin, Germany

Tel.: +49 30 206176-74

Fax: +49 30 206176-71

E-Mail: [annette.hoxtell@hwtk.de](mailto:annette.hoxtell@hwtk.de)

URL: <https://www.hwtk.de/prof-dr-annette-hoxtell/>

## Zitation

Hoxtell, Annette (2019). Eine Skizze zur Automatisierung qualitativer Inhaltsanalyse [23 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(3), Art. 15, <http://dx.doi.org/10.17169/fqs-20.3.3340>.